DOCUMENT RESUME

ED 452 822 IR 020 659

AUTHOR	Baker, Eva L.
TITLE	Technology: How Do We Know It Works?
PUB DATE	1999-00-00
NOTE	7p.; In: The Secretary's Conference on Educational Technology, 1999: Evaluating the Effectiveness of Technology. [Proceedings] (Washington, D.C., July 12-13, 1999); see IR 020 654.
AVAILABLE FROM	For full text: http://www.ed.gov/technology/1999/whitepapers.
PUB TYPE	Guides - Non-Classroom (055) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Computer Assisted Instruction; *Computer Uses in Education; Educational Administration; *Educational Development; Educational Planning; *Educational Technology; Elementary Secondary Education; Evaluation Methods; *Program Evaluation; Systems Approach; Technology Integration
IDENTIFIERS	*Technology Implementation

ABSTRACT

Does educational technology work? Does its application in classrooms help children learn? Can teachers improve their own understanding and practice through technology? The answer is yes; however in a period of widespread concern about educational quality, teachers, parents, policymakers, and taxpayers deserve answers that go beyond fervent beliefs and jaunty assertions. To sustain and support the growth of high-quality technology in schools, everyone has to learn to be more aware of what standards of documentation are useful. This paper presents a set of ideas and guidelines to consider, ending with how technology itself can aid in the testing and evaluation process. The ideas presented are intended to help educators and others invested in education think systematically about what they are doing and how to capture and document accomplishments. Highlights include: "Technology for What?"; "Technology Innovations: The How"; "Who Benefits?"; "Other Evaluation Considerations"; "Measures of Outcome and Impact"; and "Technological Supports for Evaluation." (AEF)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 452 822

Technology: How Do We Know It Works?

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

By: Eva L. Baker

IR020659

BEST COPY AVAILABLE

Technology: How Do We Know It Works?

Eva L. Baker
National Center for Research on Evaluation, Standards and Student Testing (CRESST)
UCLA

Does educational technology work? Does its application in classrooms help children learn? Can teachers improve their own understanding and practice through technology? Yes! Yes!! And yes!!!

In a period of widespread concern about educational quality, teachers, parents, policymakers, and taxpayers deserve answers that go beyond fervent beliefs and jaunty assertions. They need evidence in order to calm their doubts, justify their expenditures, and strengthen their confidence in what we do. Because we have developed the most sophisticated evaluation methods in the world, we should be able to document strengths and identify shortfalls of technology-based learning systems.

Presented here are a brief set of ideas and guidelines for you to consider, ending with how technology itself can aid in the testing and evaluation process. Let's start with the core notion that evaluation should be planned at the beginning of an innovation rather than tacked on at its end. Evaluation is a planning tool as well as a way to systematically collect and interpret findings and document impact. Scholars (Baker &Alkin, 1973, Scriven, 1967) have divided evaluation into two types: formative evaluation-where information focuses on program improvement; and summative evaluation-where information is used to make a decision among options or to certify the effectiveness of a program. In reality, all evaluation is now both summative and formative: Data help designers and users to improve practice (because nothing works right the first or second time) and also give information about whether the innovation is sufficiently promising to continue investing in. Technical standards for the conduct of evaluation have been produced (AERA, APA, & NCME, *Standards for Educational and Psychological Testing*, 1985). The degree to which evaluations demand many of these concerns depends on where the innovation is going and who has to be convinced. Who are the main consumers of information-the teachers and students in the innovation, its funders, policymakers? Does the evaluation need the blessing of an external evaluator or consultant, to give an arm's length picture of the process, or are you comfortable about building and improving your own systems? Decisions on this score may help you decide whether TO solicit external help or do it yourself. If the former, the guidelines may help you design the kind of request for proposal you want and the kind of standards for work you will accept from a subcontractor. In either case, the ideas below are intended to help you think systematically about what you're doing and how to capture and document accomplishments.

Technology for What?

What is the technology intended to do? Tom Glennan distinguishes between technology-pull and

technology-push (Glennan & Melmed, 1996). Goals for classroom technology can focus on learning the use of tools to assist in other areas of learning-for instance, using search engines, e-mail, databases, spreadsheets and word processing to find, analyze, represent, and produce documents and other products to display learning. This type of learning may be related to standards set for the school or the state's children to meet. The focus is on using technology to meet requirements. The requirements pull the technology to them.

A second set of goals may be to use technology power to address new goals that cannot be met in any other way. These could involve the designing of complex simulations, or the collaborative interaction on projects with scientists, other experts, and other students across the nation and the globe. In this case, the technology itself pushes users to new goals and new options.

The object of a third set of goals is to use technology more efficiently to deliver instructional opportunities that match the background and pace of the learners. Such uses typically involve integrated programs where students are helped to acquire specific knowledge and skills.

There are also technologies that focus on the management of classrooms by teachers, but for the moment, let us address the evaluation of students' learning. It is absolutely critical for evaluation to determine the degree of emphasis among the three kinds of goals identified above, to be clear about them, to communicate them to all collaborators, including students, and if revisions occur, to be open about how goals have changed.

Technology Innovations: The How

In every evaluation we have conducted, the road is rocky at the beginning. And whether the evaluation is well funded or operating on a shoestring, hardware and software may not arrive when expected, infrastructure may be delayed or wrong and in need of adjustment, technical assistance may be not fashioned exactly to meet the users' emerging needs. So expect this small amount of chaos.

A good evaluation considers, in addition to goals, who the key participants are-administrators, teachers, parents, students, software providers, consultants-and whose roles are key at what points. Remember also to note how decisions are made to adjust the program, whether they are explicit, and how to keep track of them. This part of evaluation is just good planning.

Implementation of an innovation also depends on a lot of different factors. First, perhaps, is the locus of the ideas for the work. Is it a school-based innovation led by teachers? Is it a collaborative venture involving software that needs to be customized and integrated into a curriculum for particular students or regions? Is it an externally imposed "opportunity" depending upon volunteers or incentives? How systematic is the use of the innovation over what time period? Are we talking about a neat activity that takes a week, or a long-term set of skills (such as modeling and representing data) that can be useful over the long haul and in which it takes a substantial time to develop expertise? Is the project one that emphasizes motivation? The excitement of communication with other students rather than the development of content expertise?

How much documentation about implementation is needed? The schedule and timeline of the beginning and key junctures in the innovation? The integration (or lack thereof) with regular parts of the curriculum? Training requirements and systems for teachers, students, and other participants?

Are the learning topics intended to include the full range of the curriculum? To focus on certain subjects, for instance, history? To concentrate on one or two topics within courses, like earthquakes in an earth sciences course? Is the emphasis interdisciplinary? Is the topic a matter of student choice, and if so, how is activity linked to important expectations?

What is the scope of the project? A few teachers at one school? Teams of teachers at the same grade in a part or all of a district? A statewide scale-up of computer-based curricula? Foundation-supported innovations of different characters and goals at different sites?

Who Benefits?

Which children or students (and teachers) are the key beneficiaries of the innovation? Is there specific background learning or experience that makes children particularly ready for the innovation planned-including language, computer skills or lack thereof, out-of-school experiences, content knowledge? Are the children located at a particular age range or grade level? Are they supposed to be affected over a number of days, weeks, or years? What is a fair comparison group? Others in the school? Children at other schools or sites?

Other Evaluation Considerations

An innovation also has a set of philosophic underpinnings that might need to be considered. Is emphasis placed on exploration and collaboration? On mastery and fluency? On subject matter depth or generalization to a number of topics and subjects? Each of these potential emphases, and many others, of course, may need to be evaluated.

Measures of Outcome and Impact

A few words of advice. Don't hinge the evaluation findings on who likes what. Teachers' descriptions of their "excitement" and students' enthusiasm are certainly desirable, but are probably unlikely to persuade external decision makers of the success of an innovation by itself. If that enthusiasm links to fewer absences, or more attentiveness, then the evaluation will gain power. As an overall dictum, focus first, intensely, and last on student learning. Such a concentration will refer you back to your original goals and may require a redefinition of your original intentions.

Measuring outcomes involves two main components: what you will use to provide the data, and how you will decide whether the findings are sufficiently good to warrant continuation, revision, and so on.

Types of measures include regularly administered tests, either commercial or statewide assessments. There may be special tests already available to measure students' acquisition of the particular area of focus. Often the tests and measures may need to be developed to tap into new uses to which the computer is put. These other measures may include projects, essays, and extended performances, as well as typical tests of knowledge and skills. You need to be sensitive to the fact that if you use open-ended tasks such as performance or essay examinations, you need to use clear criteria to judge performance, and performance should be validly and consistently measured among raters. You should remove, to the extent you can, the bias inherent in having teachers rate their own students or the performance of only students known to be in the technology option. Questionnaires asking about student attitude, ease of use of the applications, and suggestions for improvement from those who participated in the technology may also

be helpful.

The most frequent way that evaluators determine whether performance is good enough is by using comparisons. You can compare students in and out of the innovation (although to be certain, you should assign them randomly rather than just using intact classroom groups). You can use pretest versus posttest scores, particularly if you have comparison groups of similar students. If you use pretests and posttests, you'll probably need some external help to deal with the practice effects of the test (learning from the test), interaction effects (how the pretest may enhance the impact of the technology), and the reliability of the measure you use (the difference between pre and post, or a more sophisticated statistical analysis). All this help is readily available. You may also want to follow up students and look at their performance over time, even after they are through with the particular program of interest, in order to determine whether there are long-term effects. When you have sufficient numbers, you should disaggregate your results to see whether the innovation works better for students with certain backgrounds, particular experiences, or specific knowledge.

The major trade-off is whether you link results to the regular test (policymakers would like that) and recognize that it is generally much harder to show impact in this way than on assessments targeted toward the same content and cognitive demands as the innovation. Your local policies may be your best guide here.

Technological Supports for Evaluation

It makes most sense, of course, to use measures that optimize detection of impact for the innovation you are developing. For that reason, we advocate the use of computer-based assessments where possible and where they have sufficient technical quality, including validity and reliability evidence. CRESST has developed measures of problem solving, content understanding, knowledge representation, search strategies, collaboration, and Internet learning, for example, that can be administered by computers. Ideally, you would want to automate information about how students are engaging in their technology use to help you understand why you have obtained given results. Maybe students do best who have a slowly increasing involvement. Maybe there is a threshold that allows them to take off. Maybe their lack of background content knowledge is holding them back.

A second kind of support that CRESST has is a database manager (called the Quality School Portfolio, or QSP) that allows the user to transform databases (for instance, of district or state scores) into a local, longitudinal database for all students. Then students in the technology innovation and those in the comparison group can be sampled on various bases-background, prior subject matter grades, test scores-and the data disaggregated immediately. QSP also allows the use of locally developed outcome and attitudinal measures by providing a resource kit of measures, guidelines for their use, and scanning and analytical capability. In the end, QSP generates a report comparing groups, or a single group at multiple time points. Graphical reporting can be tailored to various audiences for the report.

Summary

To sustain and support the growth of high-quality technology in schools, everyone has to learn to be more aware of what standards of documentation are useful. Each of us can learn to interpret quality information to revise, redesign, or reconceive the ways technology can be used to help our children meet our expectations. Better that we have a hand in it.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Baker, E. L., & Alkin, M. C. (1973). Formative evaluation of instructional development. *AV Communication Review, 21*(4), 389-418. (ERIC Document Reproduction Service No. EJ 091 462)

Glennan, T. K., & Melmed, A. (1996). *Fostering the use of educational technology: Elements of a national survey* (MR-682-OSTP/ED). Santa Monica, CA: RAND.

Scriven, M. (1967). The methodology of evaluation. In R. E. Stake and others (Eds.), *Perspectives on curriculum evaluation.* AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally.

-###-

[The Technology/Content Dilemma]

[Documenting the Effects of Instructional Technology: A Fly-Over of Policy Questions]